

SPEECH RECOGNITION AND RESPONSE DEVICE

Patent Number: JP8265440
Publication date: 1996-10-11
Inventor(s): OOKAWA HIROTOSHI; TOMARI YOICHIRO; NAKATSUKA TETSUO
Applicant(s): MITSUBISHI ELECTRIC CORP
Requested Patent: ☐ JP8265440
Application Number: JP19950069403 19950328
Priority Number(s):
IPC Classification: H04M3/50; G10L3/00; G10L3/00; H04M3/42
EC Classification:
Equivalents:

Abstract

PURPOSE: To enable a user to listen to a response voice matched with the user himself (or herself) by making it possible to change the output mode of a response voice in accordance with an instruction from the user.

CONSTITUTION: This speech recognition and response device is provided with a voice response part 45 capable of executing voice response corresponding to the recognized result of a voice signal inputted by a user through a telephone line network 2, a control signal recognizing part 44 for recognizing a control signal inputted by the user through the network 2 during the output of a response from the response part 45 and a recognized response control part 46 for outputting a response output control signal to the response part 45 so that the response part 45 outputs a response based upon the mode of a response output executed up to now correspondingly to the recognized result of the recognizing part 44. When the control signal is inputted by the user during the period of a voice response, various modes for responses between the device and the user can be attained in accordance with an instruction from the user based upon stored data related to voice responses. Namely a response voice speed is changed. In addition, fast-feeding, rewinding, etc., for a response voice can be executed.

Data supplied from the esp@cenet database - 12

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平8-265440

(43)公開日 平成8年(1996)10月11日

(51)Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 M 3/50			H 0 4 M 3/50	A
G 1 0 L 3/00	5 5 1		G 1 0 L 3/00	5 5 1 A
	5 6 1			5 6 1 D
H 0 4 M 3/42			H 0 4 M 3/42	N
				P
審査請求 未請求 請求項の数14 O L (全 19 頁)				

(21)出願番号 特願平7-69403

(22)出願日 平成7年(1995)3月28日

(71)出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目2番3号

(72)発明者 大川 裕利

鎌倉市上町屋325番地 三菱電機株式会社

情報システム製作所内

(72)発明者 泊 陽一郎

鎌倉市上町屋325番地 三菱電機株式会社

情報システム製作所内

(72)発明者 中塚 哲夫

鎌倉市上町屋325番地 三菱電機株式会社

情報システム製作所内

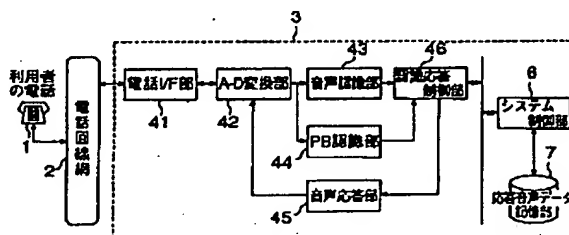
(74)代理人 弁理士 高田 守 (外4名)

(54)【発明の名称】 音声認識応答装置

(57)【要約】

【目的】 電話を介して入力された音声信号及び制御信号を認識し、認識結果をもとに応答する音声認識応答装置において、利用者の使い勝手の向上、適正な対話の確保、さらにオペレータ処理の負荷の低減を図る。

【構成】 応答音声出力中に利用者が入力した制御信号を認識し、その認識結果に基づき応答音声の早送り、巻き戻し等を行うための認識応答制御部を設けた。認識が不能になれば、利用者の指示に基づきオペレータに処理が移り、対話が正常になったとオペレータが判断すれば、再び利用者は装置と対話するよう制御部を設けた。



【特許請求の範囲】

【請求項1】 利用者により電話回線を通して入力された音声信号を認識する音声認識部、
上記音声認識部の認識結果に対応した音声応答を行う音声応答部、

上記音声応答部の応答出力中に、利用者により上記電話回線を通じて入力された制御信号を認識する制御信号認識部、

上記制御信号認識部の認識結果に対応して、音声応答部が、それ迄に実行していた応答出力の態様と態様の応答出力を行うよう応答出力制御信号を、上記音声応答部に出力する認識応答制御部、を備えた音声認識応答装置。

【請求項2】 音声応答部による音声応答に関するデータを記憶する記憶部、

上記認識応答制御部と関連して設けられ、上記記憶部によるデータの書込み／読出し動作を制御する制御部、を更に備えたことを特徴とする請求項1の音声認識応答装置。

【請求項3】 応答音声の出力中に、利用者の電話機からのプッシュボタン操作に基づく指示に基づき、認識応答制御部は応答音声の速度を選択するための応答出力制御信号を音声応答部に出力することを特徴とする請求項2の音声認識応答装置。

【請求項4】 応答音声の出力中に、利用者の電話機からのプッシュボタン操作に基づく指示に基づき、認識応答制御部は応答音声の早送り、後戻りを選択するための応答出力制御信号を音声応答部に出力することを特徴とする請求項2の音声認識応答装置。

【請求項5】 装置からの問合わせを応答音声として出力し、利用者は電話機から音声で応え、応答音声出力中に、利用者は電話機からのプッシュボタン操作に基づく指示に基づき、記憶応答制御部に出力中の応答音声をスキップするための応答出力制御信号を音声応答部に出力することを特徴とする請求項2の音声認識応答装置。

【請求項6】 問合わせと利用者の音声入力について、応答音声出力中に、利用者は電話機からのプッシュボタン操作に基づく指示に基づき、出力中の応答音声を中断し一つ前の問合わせに対応した応答音声を出力させるよう認識応答制御部は応答出力制御信号を音声応答部に出力することを特徴とする請求項5の音声認識応答装置。

【請求項7】 音声認識部で電話機から入力された音声認識できなかった場合、認識応答制御部は、入力された音声の後に予め定められた音声を付加した音声を応答音声として出力するよう応答出力制御信号を音声応答部に出力することを特徴とする請求項2の音声認識応答装置。

【請求項8】 音声認識部で電話機から入力された音声認識できなかった場合、音声認識部で検出した音声の始端と終端で挟まれた音声の後に予め定められた音声を付加した音声を応答音声として出力するよう認識応答制

御部は応答出力制御信号を音声応答部に出力することを特徴とする請求項2の音声認識応答装置。

【請求項9】 音声認識部が、聞き違い易い複数の音声認識結果を出した場合、応答音声で複数の音声認識結果を電話から利用者に知らせ、利用者からの応答により、利用者の入力音声を確認することを特徴とする請求項2の音声認識応答装置。

【請求項10】 入力音声の確認を、数字音声の入力により行うことを特徴とする請求項9の音声認識応答装置。

【請求項11】 予め聞き違い易い音声をグループ化したデータとして保持しておき、音声認識部が、聞き違い易い音声の一つを認識結果として出した場合、応答音声でその音声の属するグループを、電話機から利用者に知らせ、利用者からの応答により、利用者の入力音声を確認することを特徴とする請求項2の音声認識応答装置。

【請求項12】 利用者の入力音声の確認を、数字音声の入力により行うことを特徴とする請求項11の音声認識応答装置。

【請求項13】 応答音声の大きさ、音質として男性の声、女性の声の指定、結果、応答音声の速さなどの値を記憶し、次回同一の利用者の利用があったとき、記憶結果をもとに、応答音声を出力することを特徴とする請求項第2項の音声認識応答装置。

【請求項14】 利用者により電話回線を通して入力された音声信号を認識する音声認識部と、上記音声認識部の認識結果に対応した音声応答を行う音声応答部と、
上記音声応答部の応答出力中に、利用者により上記電話回線を通じて入力された制御信号を認識する制御信号認識部と、

上記制御信号認識部の認識結果に対応して音声応答部が、それ迄に実行していた応答出力の態様と態様の応答出力を行うよう応答出力制御信号を、上記音声応答部に出力する認識応答制御部とを有する応答手段と、
利用者によるオペレータ呼出指示に基づく利用者とオペレータ側との上記応答手段を介した応答交信とオペレータ側からの指示に基づく利用者と上記応答手段との、応答交信とを制御する制御部とを、
備えた音声認識応答装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、電話機などから電話回線を通して入力された音声及び制御信号を認識し、認識結果をもとに応答する音声認識応答装置（以下「認識応答装置」と呼ぶ）に関するものである。利用者は、電話機などから、認識応答装置に電話し、認識応答装置の応答音声に応じて、必要項目を音声入力することで、商品の発注や問合わせなどを行う。

【0002】

【従来の技術】認識応答装置の一つの問題点として、装

置に慣れた利用者也毎回同じ案内音声を開かなければならないことや、一方で聞き逃した音声を再度聞くためにもう一度最初から聞かなければならないこと、装置の応答音声に対して一つ一つ応えていかなければ次の応答音声を聞けないこと等、いわゆる使い勝手の悪さがある。

(問題点 1)

【0003】使い勝手の悪さを解消する従来技術として、例えば、図 25 及び図 26 に示す特公昭 62-40577 では、特定語を認識する特定語認識回路を設け、特定語が認識された場合には特定語が音声入力された時点の語群から再びメッセージを出力する出力回路を設け、迅速な聞き返しができるようにしている。

【0004】この従来技術は、図 25 に示すように、音声認識部 25-1 と、音声応答部 25-3 及びこれらを制御する会話制御部 25-2 を備えている。処理の流れは、音声応答部 25-3 から、応答音声出力されている時に、利用者が、「えっ」、「なんですって」等と発声した時、図 26 に示すように、まずステップ S 26-1 で音声の入力を音声認識部 25-1 で検知し、その結果ステップ S 26-2 で会話制御部 25-2 から、音声応答部 25-3 に音声合成の中断が指示される。ステップ S 26-3 で音声認識部 25-1 により入力音声を認識し、その結果、「えっ」、「なんですって」等の特定語であれば、ステップ S 26-4 から分岐したステップ S 26-5 で、文節、単語単位で音声合成を再合成をするよう音声応答部 25-3 に指示があり、音声と言い直される。一方、特定語でなければ、ステップ S 26-6 で音声合成の中断を解除し、引き続き応答音声を出力する。しかし、この技術では、特定語に誤認識があった場合に、対応できないという欠点がある。(問題点 2)

【0005】認識応答装置のもう一つの問題点は、音声認識技術によって認識率は向上しているものの、必ず誤認識を伴うものであることに起因する。誤認識によって、利用者の意図とは異なる言葉が入力され、利用者の意図しないデータが装置から応答音声として出力されたり、意図しない方向へ利用者と装置の対話が進んだりすることがあり、これを最小限にする必要がある。(問題点 3)

【0006】誤認識によって、利用者の意図しない方向へ対話が進んだ場合の救済手段の従来技術として、例えば図 27 及び図 28 に示す特公平 3-88466 では、音声認識機能で認識不能の場合はオペレータに接続を切り替えることとしている。

【0007】この従来技術では、図 27 に示すように、音声認識音声合成装置 27-5、アナウンスメント用データベース 27-6、パケットデータ網または専用線 27-7、番号案内情報データベース 27-8、及びオペレータ端末 27-9 を備え、これらの接続を切り替えるスイッチ (SW) 27-3 及び、その切り替えを制御する CPU 27-4 を備えている。

【0008】処理の流れは、図 28 に示すように、まずステップ S 28-1 で利用者からかかってきた電話の接続があり、次にステップ S 28-2 でアナウンスメント用データベース 27-6 から引き出した応答音声を、音声認識音声合成装置 27-5 から出力する。ステップ S 28-3 で利用者からの入力音声を音声認識合成装置 27-5 で認識し、28-4 で音声認識処理が正常であれば、ステップ S 28-5 で終了かどうかを判断し、利用者の入力項目があれば、ステップ S 28-2 で次の応答音声を出力し、ステップ S 28-3 入力音声の認識というシーケンスを繰り返す。ステップ S 28-5 で利用者の入力項目がなく、終了であれば、ステップ S 28-6 パケットデータ網または専用線 27-7 を通じて、番号案内情報データベース 27-8 からデータを取り出し、27-5 音声認識音声合成装置から、応答音声を出力して、処理を終わる。

【0009】一方、ステップ S 28-4 で音声認識が正常でないと判断されれば、ステップ S 28-7 で、CPU 27-4 は SW 27-3 を、オペレータ 27-9 に接続する。ステップ S 28-8 で以降はオペレータ 27-9 が利用者に対話し、処理を終了する。以上のようにこの技術では、誤認識の発声で音声認識が正常にできなくなった場合、オペレータへの処理が切り替えられ、利用者とオペレータの対話は処理の終了まで続く。一度の音声認識の異常で、オペレータに処理が切り替えられると、以後はオペレータにまかされる。特に、対話の初期で音声認識が異常になると、オペレータの負荷は、最初からオペレータが利用者に対話している場合と変わらず、音声認識の機能が充分活用されず、依然としてオペレータの負荷が大きいという問題点がある。(問題点 4)

【0010】

【発明が解決しようとする課題】従来技術によると、認識応答装置の使い勝手の向上や誤認識が発生した場合の対応が示されているが、まだ不十分で以下の問題点があった。

【0011】問題点 1、聞き逃しに対する再出力、既に知っている応答音声に対する聞き飛ばしなど、利用者の柔軟な要求に対応できない。

【0012】問題点 2、従来の技術によると、聞き逃した際の応答音声の再出力について、特定語の認識を用いた例があるが、特定語そのものの誤認定には充分には対応できない。

【0013】問題点 3、誤認識を低減し、利用者の意図しない方向へ対話が進むのを防ぐには、音声認識技術の向上の他に、利用者と認識応答装置が協力して、入力音声を確定させることが必要であり、この点につき不十分であった。

【0014】問題点 4、従来の技術で、認識不能になった場合オペレータに接続を切り替える例があるが、これ

5

だけでは、誤認識が起こり一度処理がオペレータに切り替えられると、後はオペレータまかせになるため、特に利用者からの入力項目が多い場合に、最初の方でオペレータに切り替わると、以降のオペレータの負荷が増し、省力化を目的とした認識応答装置本来の目的を果たすことが困難である。

【0015】この発明は、上記のような問題点を解消するためになされたもので、以下に述べる種々の目的のうち少なくともいずれかを達成するようにしたものである。

1. 利用者の柔軟な要求に対応できる。
2. 聞き逃しに対して充分に対応できる。
3. 誤認識を低減し、利用者の意図しない方向へ認識が進むのを防ぐことができる。
4. オペレータの介入の必要性を低減し、オペレータの負荷を軽減することができる。

【0016】

【課題を解決するための手段】上記目的を達成するための本発明の音声認識応答装置では、利用者により電話回線を通して入力された音声信号の認識結果に対応した音声応答を行う音声応答部と、音声応答部の応答出力中に、利用者により上記電話回線を通じて入力された制御信号を認識する制御信号認識部と、制御信号認識部の認識結果に対応して音声応答部が、それまでに実行していた応答出力の態様と態様の応答出力を行うよう応答出力制御信号を、音声応答部に出力する認識応答制御部とを、設けたものである。

【0017】また本発明では、更に、音声応答部による音声応答に関するデータを記憶する記憶部と、記憶部によるデータの書込み／読出し動作を制御する制御部とを設けたものである。

【0018】また本発明では、認識応答制御部から音声応答部に対し、適宜以下のようないずれかの応答出力制御信号を出力するようにしたものである。

1. 応答音声速度を選択するための応答出力制御信号。
2. 応答音声の早送り、後戻りを選択するための応答出力制御信号。
3. 出力中の応答音声をスキップするための応答出力制御信号。

【0019】また本発明では、装置から利用者の入力を促す応答音声について、プッシュボタン操作により、出力中の応答音声を中断するようにしたものである。

【0020】また本発明では、音声認識部で入力された音声を認識できなかった場合、入力した音声の後に固定の音声を付加した音声を示す応答音声として出力するようにしたものである。

【0021】また本発明では、音声認識部で入力された音声を認識できなかった場合、音声認識部で入力した音声を録音し、音声認識部で検出した音声の始端及び終端

6

に挟まれた音声の後に、予め定められた音声を付けて応答音声を出力するようにしたものである。

【0022】また本発明では、音声認識部で、聞き違い複数の認識結果を出した場合、システム制御部にその結果を伝え、システム制御部では、応答音声として、複数の認識結果を利用者に伝え、利用者からの応答により入力音声を確定するようにしたものである。

【0023】また本発明では、予め聞き違い易い音声をグループ化し、類似語グループとし、音声認識部が、類似語グループに属する音声の一つを結果として出した場合、システム制御部では、類似語グループに属する音声を利用者に伝え、利用者からの応答により入力音声を確定するようにしたものである。

【0024】また本発明では、応答音声の大きさ、音質として男性の声または女性の声の指定結果、応答音声の速さを記憶保持しておき、次回同一の利用者からの利用があったとき、記憶結果をもとに応答音声を出力するようにしたものである。

【0025】また本発明では、利用者によるオペレータ呼出指示に基く利用者とオペレータ側との応答交信とオペレータ側からの指示に基く、利用者と上記応答装置との応答交信とを制御する制御部を設けたものである。

【0026】

【作用】上記構成により本発明によれば、音声応答中に利用者により制御信号が入力されると、応答出力の態様が変化する。

【0027】音声応答に関する記憶データに基き、装置と利用者との応答につき利用者からの指示により、以下に示すように、種々の態様が実現される。

1. 応答音声速度が変化する。
2. 応答音声の早送り、後戻りがなされる。
3. 出力中の応答音声をスキップする。

【0028】また本発明によれば、音声認識部で、聞き違い複数の認識結果を出した場合、その入力音声の確定を利用者の応答に基き行われ、聞き違い易い音声容易に確定される。

【0029】また本発明によれば、予め聞き違い易い音声を類似語グループとしてグループ化しておき、グループの中の一つが認識結果としてでた場合、利用者の応答に基き入力音声確定され、聞き違い易い音声に起因する誤認識が発生した場合にも、入力音声容易に確定される。

【0030】前回の応答音声の大きさ、速さ、温室等を次回同一の利用者の利用があった場合に、応答音声を前回の態様にて出力される。

【0031】また本発明によれば、利用者の指示に基き、利用者とオペレータとの対話応答交信がなされ、オペレータの指示に基き、利用者と装置との対話応答交信がなされ、オペレータが利用者対話する機会が減少し、オペレータの負荷が軽減する。

【0032】

【実施例】

実施例 1. 図 1 は、本発明の第 1 実施例に係る音声認識応答装置の全体構成を示すブロック図である。

【0033】図 1 で、1 は利用者の電話、2 は電話回線網、3 は音声認識応答装置、4 1 は音声認識応答装置が電話回線網につながるための電話 I/F 部、4 2 は電話 I/F 部からきたアナログ信号をディジタル信号に変換したり、ディジタル信号をアナログ信号に変換して電話 I/F 部に送る A-D 変換部、4 3 は A-D 変換部からきたディジタル信号を音声認識する音声認識部、4 4 は A-D 変換部からきたディジタル信号をプッシュボタン信号としてプッシュボタン認識する PB 認識部、4 5 は A-D 変換部に応答音声データを出力する応答音声部、4 6 は音声認識部、PB 認識部、応答音声部を制御する認識応答制御部、6 は音声認識応答装置全体を制御するシステム制御部、7 は応答音声データが圧縮されて記憶されている応答音声データ記憶部である。

【0034】図 2 は、利用者電話のプッシュボタンが押下された時の処理の流れを示すフロー図である。図 3 は、利用者電話のプッシュボタンを示したものである。

【0035】図 4 は、図 2 の利用者電話のプッシュボタンについて、応答音声出力中にボタンプッシュすることで、働く機能を示したものである。

【0036】図 5 は、利用者と認識応答装置の対話における装置側から出される質問項目と、それを構成する応答音声ファイルの番号の対応を示したものである。

【0037】次に動作について説明する。図 2 のフロー図において、利用者がステップ S 2-1 において電話接続した後、ステップ S 2-2 において、応答音声の出力が行われる。応答音声の速度を速くしたい時、図 4 に示すようにプッシュボタンの「1」を押下する。ステップ S 2-3 でプッシュボタンの入力があるかどうか判断され、あった場合はステップ S 2-4 でプッシュボタン認識が行われる。そして、それが「1」であれば、ステップ S 2-5 で応答音声の速度を速くする処理が行われる。

【0038】以上を図 1 で説明すると、利用者の電話 1 から入力されたプッシュボタン信号は、電話回線網 2、電話 I/F 部 4 1、A-D 変換部 4 2 を通じて、PB 認識部 4 4 に伝えられる。PB 認識部 4 4 では、プッシュボタン信号を認識し、プッシュボタンにより、「1」が入力されたことを、認識応答制御部 4 6 に伝える。認識応答制御部 4 6 は、「速く」が指示された場合の応答音声の速度係数を音声応答部 4 5 に伝える。次の応答音声出力から、音声応答部 4 5 は、応答音声データファイルを分解する処理の中で、速度係数に応じて応答音声の速度を速くする処理を行って、A-D 変換部 4 2 に送り、応答音声出力される。

【0039】応答音声の速度を速くする方法としては、例えば以下のようなものが考えられる。応答音声の波形

を解析して、周期を検出する。そして、その周期を元に、周期が変わらないようにして、音声波形全体が指定した速度係数の逆数倍の長さになるように、データを抜いてやる。例えば、速度係数を 1.25 倍にする場合、長さはその逆数なので 0.8 倍にしてやればよい。このためには、5 周期ごとに 1 周期分、データを抜くようにする。

【0040】利用者は、応答音声を普通速度に戻したい時、図 4 に示すプッシュボタンの「2」を押下する。処理の流れは図 2 で、ステップ S 2-3 でプッシュボタンの入力があるかどうか判断され、あった場合はステップ S 2-4 でプッシュボタン認識が行われる。そして、それが「2」であれば、ステップ S 2-6 で応答音声を普通速度にする処理が行われる。

【0041】以上を図 1 で説明すると、利用者の電話 1 から入力されたプッシュボタン信号は、電話回線網 2、電話 I/F 部 4 1、A-D 変換部 4 2 を通じて、PB 認識部 4 4 に伝えられる。PB 認識部 4 4 では、プッシュボタン信号を認識し、プッシュボタンにより、「2」が入力されたことを、認識応答制御部 4 6 に伝える。認識応答制御部 4 6 は、「普通速度」の応答音声の速度係数 1.0 を音声応答部 4 5 に伝える。次の応答音声出力から、音声応答部 4 5 は、応答音声の長さを変化させず、そのまま A-D 変換部 4 2 に送り、応答音声出力される。

【0042】利用者は、応答音声の速度を遅くしたい時、図 4 に示すプッシュボタンの「3」を押下する。処理の流れは図 2 で、ステップ S 2-3 でプッシュボタンの入力があるかどうか判断され、あった場合はステップ S 2-4 でプッシュボタン認識が行われる。そして、それが「3」であれば、ステップ S 2-7 で応答音声を遅くする処理が行われる。

【0043】以上を図 1 で説明すると、利用者の電話 1 から入力されたプッシュボタン信号は、電話回線網 2、電話 I/F 部 4 1、A-D 変換部 4 2 を通じて、PB 認識部 4 4 に伝えられる。PB 認識部 4 4 では、プッシュボタン信号を認識し、プッシュボタンにより、「3」が入力されたことを、認識応答制御部 4 6 に伝える。認識応答制御部 4 6 は、「遅く」が指示された場合の応答音声の速度係数を音声応答部 4 5 に伝える。次の応答音声出力から、音声応答部 4 5 は、応答音声データファイルを分解する処理の中で、速度係数に応じて応答音声の速度を遅くする処理を行って A-D 変換部 4 2 に送り、応答音声出力される。

【0044】応答音声の速度を遅くする方法としては、例えば以下のようなものが考えられる。応答音声の波形を解析して、周期を検出する。そして、その周期を元に、周期が変わらないようにして、音声波形全体が指定した速度係数の逆数倍の長さになるように、同じ音声波形をコピーして挿入する。例えば、速度係数を 0.8 倍

にする場合、長さはその逆数なので1.25倍にしてやればよい。このためには、4周期ごとに1周期分、同じ音声波形をコピーして挿入する。

【0045】利用者は、応答音声現在の速度よりも速くしたい時、図4に示すプッシュボタンの「4」を押下する。処理の流れは図2で、ステップS2-3でプッシュボタンの入力があるかどうか判断され、あった場合はステップS2-4でプッシュボタン認識が行われる。そして、それが「4」であれば、ステップS2-8で応答音声を徐々に速くする処理が行われる。

【0046】以上を図1で説明すると、利用者の電話1から入力されたプッシュボタン信号は、電話回線網2、電話1/F部41、A-D変換部42を通じて、PB認識部44に伝えられる。PB認識部44では、プッシュボタン信号を認識し、プッシュボタンにより、「4」が入力されたことを、認識応答制御部46に伝える。認識応答制御部46は、「より速く」が指示された場合の速度係数の増加分だけ現在の応答音声の速度係数に加算し、それを音声応答部45に伝える。次の応答音声出力から、音声応答部45は、応答音声データファイルを分解する処理の中で、速度係数が1以上ならばそれに応じて一定の間隔でデータを抜くことにより、応答音声短くなるようにして、また速度係数が1以下ならばそれに応じて一定の間隔で同じデータを挿入することにより、応答音声長くなるようにして、A-D変換部42に送り、応答音声出力される。プッシュボタンが繰り返し押下された時は、その回数分だけ、速度係数を加算するので、応答音声はより速くなる。

【0047】利用者は、応答音声現在の速度よりも遅くしたい時、図4に示すプッシュボタンの「6」を押下する。処理の流れは図2で、ステップS2-3でプッシュボタンの入力があるかどうか判断され、あった場合はステップS2-4でプッシュボタン認識が行われる。そして、それが「6」であれば、ステップS2-9で応答音声を徐々に遅くする処理が行われる。

【0048】以上を図1で説明すると、利用者の電話1から入力されたプッシュボタン信号は、電話回線網2、電話1/F部41、A-D変換部42を通じて、PB認識部44に伝えられる。PB認識部44では、プッシュボタン信号を認識し、プッシュボタンにより、「6」が入力されたことを、認識応答制御部46に伝える。認識応答制御部46は、「より遅く」が指示された場合の速度係数の減少分だけ現在の応答音声の速度係数に減算し、それを音声応答部45に伝える。次の応答音声出力から、音声応答部45は、応答音声データファイルを分解する処理の中で、速度係数が1以上ならばそれに応じて一定の間隔でデータを抜くことにより、応答音声短くなるようにして、また速度係数が1以下ならばそれに応じて一定の間隔で同じデータを挿入することにより、応答音声長くなるようにして、A-D変換部42に送

り、応答音声出力される。プッシュボタンが繰り返し押下された時は、その回数分だけ、速度係数を減算するので、応答音声はより遅くなる。

【0049】応答音声は、1文あるいは句読点で切られる単位ごとに応答音声データ記憶部7の応答音声データファイルに保存されており、ファイル番号で管理されている。例えば、図5に示すように、認識応答装置から、利用者に向けて問合わせの応答音声Q1、Q2、Q3、Q4とあったとして、Q1は、応答音声ファイル番号が101、102、103の応答音声データファイルからなっており、これらのファイルが連続して応答音声出力されることで、Q1の応答音声成り立っている。

【0050】利用者は、応答音声早送りしたい時、図4に示すプッシュボタンの「7」を押下する。例えばQ2の応答音声203を出力中にプッシュボタン「7」が押下されると、図2のステップS2-3でプッシュボタンの入力があるかどうか判断され、あった場合はステップS2-4でプッシュボタン認識が行われる。そして、それが「7」であれば、ステップS2-10で応答音声の早送りする処理が行われる。図6に通常の出力と、途中でプッシュボタン「7」の入力があった場合の違いを示す。プッシュボタン「7」の入力によって、出力中の応答音声203を中断し、次に用意された応答音声204が出力される。

【0051】図1で、利用者の電話1から入力されたプッシュボタン信号は、電話回線網2、電話1/F部41、A-D変換部42を通じて、PB認識部44に伝えられる。PB認識部44では、プッシュボタン信号を認識し、プッシュボタンにより「7」が入力されたことを、認識応答制御部46を通じて、システム制御部6に伝える。認識応答制御部46は音声応答部45を制御することで、出力中の応答音声203を中断する。システム制御部6は、次の応答音声データ204を応答音声データ記憶部7から引き出し、認識応答制御部46を通じて、音声応答部45に送り、A-D変換部42を通じて応答音声出力される。プッシュボタン「7」が繰り返し押された時は、その回数分だけ先の応答音声データを引き出して出力する。

【0052】利用者は、応答音声を巻き戻したい時、図4に示すプッシュボタンの「9」を押下する。例えばQ2の応答音声203を出力中にプッシュボタン「9」が押下されると、図2のステップS2-3でプッシュボタンの入力があるかどうか判断され、あった場合はステップS2-4でプッシュボタン認識が行われる。そして、それが「9」であれば、ステップS2-11で応答音声を巻き戻す処理が行われる。図7に通常の出力と、途中でプッシュボタン「9」の入力があった場合の違いを示す。プッシュボタン「9」の入力によって、現在出力中の応答音声を中断し、その前に出力した応答音声202の再出力を行う。

【0053】図1で、利用者の電話1から入力されたプッシュボタン信号は、電話回線網2、電話I/F部41、A-D変換部42を通じて、PB認識部44に伝えられる。PB認識部44では、プッシュボタン信号を認識し、プッシュボタンにより「9」が入力されたことを、認識応答制御部46を通じて、システム制御部6に伝える。認識応答制御部46は音声応答部45を制御することで、出力中の応答音声203を中断する。システム制御部6は、その前の応答音声データ202を応答音声データ記憶部7から引き出し、認識応答制御部46を通じて、音声応答部45に送り、A-D変換部42を通じて応答音声203が出力される。プッシュボタン「9」が繰り返し押された時は、その回数分だけ先の応答音声データを引き出して出力する。

【0054】利用者は、現在対話中の認識応答装置からの問い合わせ項目をスキップして次の問い合わせ項目に進みたい時、図4に示すプッシュボタンの「*」を押下する。例えばQ2の応答音声203を出力中にプッシュボタン「*」が押下されると、図2のステップS2-3でプッシュボタンの入力があるかどうか判断され、あった場合はステップS2-4でプッシュボタン認識が行われる。そして、それが「*」であれば、ステップS2-12で問い合わせ項目をスキップする処理が行われる。図8に通常の出力と、途中でプッシュボタン「*」の入力があった場合の違いを示す。プッシュボタン「*」の入力によって、出力中の応答音声を中断し、問い合わせ項目のスキップが行われ応答音声301の出力を行う。

【0055】図1で、利用者の電話1から入力されたプッシュボタン信号は、電話回線網2、電話I/F部41、A-D変換部42を通じて、PB認識部44に伝えられる。PB認識部44では、プッシュボタン信号を認識し、プッシュボタンにより「*」が入力されたことを、認識応答制御部46を通じて、システム制御部6に伝える。認識応答制御部46は音声応答部45を制御することで、出力中の応答音声203を中断する。システム制御部6は、その次の質問項目3の応答音声データ301を応答音声データ記憶部7から引き出し、認識応答制御部46を通じて、音声応答部45に送り、A-D変換部42を通じて応答音声203が出力される。この時、スキップされた問い合わせ項目Q2は認識応答制御部46に記憶され、他の問い合わせ項目の対話が終わった後で、未回答項目に回答することができるようになっている。

【0056】利用者は、現在対話中の問い合わせ項目を中断して一つ前の問い合わせ項目に戻りたい時、図4に示すようにプッシュボタンの「#」を押下する。例えばQ2の応答音声203を出力中にプッシュボタン「#」が押下されると、図2のステップS2-3でプッシュボタンの入力があるかどうか判断され、あった場合はステップS2-4でプッシュボタン認識が行われる。そして、それが「#」であれば、ステップS2-13で問い合わせ

項目を後戻りする処理が行われる。図9に通常の出力と、途中でプッシュボタン「#」の入力があった場合の違いを示す。プッシュボタン「#」の入力によって、出力中の応答音声203を中断し、一つ前の問い合わせ項目Q1の応答音声101の出力を行う。

【0057】図1で、利用者の電話1から入力されたプッシュボタン信号は、電話回線網2、電話I/F部41、A-D変換部42を通じて、PB認識部44に伝えられる。PB認識部44では、プッシュボタン信号を認識し、プッシュボタンにより「#」が入力されたことを、認識応答制御部46を通じて、システム制御部6に伝える。認識応答制御部46は音声応答部45を制御することで、出力中の応答音声203を中断する。システム制御部6は、一つ前の問い合わせ項目1の応答音声データ101を応答音声データ記憶部7から引き出し、認識応答制御部46を通じて、音声応答部45に送り、A-D変換部42を通じて応答音声203が出力される。

【0058】利用者が、プッシュボタンの押下により、設定した応答音声の速度、また別のプッシュボタン操作により設定した応答音声の大きさや、男性の声、女性の声などの応答音声に関して利用者が設定した値で、利用者と認識応答装置の全ての対話が終了し、電話が切られる直前の値を、システム制御部6は、図2のステップS2-17において、応答音声データ記憶部7に保存させる。次回同一の利用者から認識応答装置の利用があった時、最初の値として設定して対話を開始する。

【0059】ここで、実施例1に関する各動作を以下要約して述べる。

【0060】繰り返し押すプッシュボタン操作により、応答音声の速度を利用者の自由なレベルで速くしたり、遅くしたりする設定ができる。

【0061】応答音声データは文、文節、単語等の一定のまとまりが応答音声ファイルとして、システム制御部から応答音声部に送られ、その応答音声ファイルごとに連続したフレーム番号がつけられている。繰り返し押すプッシュボタン操作により、押下した回数に応じて、フレーム番号を元に応答音声の早送り、後戻りができる。

【0062】装置から利用者の入力を促す応答音声について、プッシュボタン操作により、出力中の応答音声を中断し、次に用意された応答音声を出力する。

【0063】装置から利用者の入力を促す応答音声について、プッシュボタン操作により、出力中の応答音声を中断し、その前に出力した応答音声を再出力する。

【0064】音声認識部で入力された音声を認識できなかった場合、音声認識部で入力した音声を録音し、録音した結果をシステム制御部に送り、システム制御部で、録音した音声を一つの応答音声ファイルとし、そのファイルの後に、それが認識できなかった旨のメッセージを示す応答音声ファイルを付けて応答音声をつくり、応答音声部から出力する。

13

【0065】音声認識部で入力された音声を確認できなかった場合、音声認識部で入力した音声を録音し、音声認識部で検出した音声の始端及び終端に挟まれた音声をシステム制御部に送り、システム制御部で、送られた音声の一つの応答音声ファイルとし、そのファイルの後に、装置では出力したように聞こえ、認識できなかった旨のメッセージを示す応答音声ファイルを付けて応答音声をつくり、応答音声部から出力する。また本発明では、プッシュボタン操作で設定した応答音声の大きさ、応答音声の速度、音質として男性の声または女性

【0066】実施例 2。図 10 は、本発明の第 2 実施例に係る認識応答装置の全体構成を示すブロック図である。図 10 で、図 1 と同じ番号は同じものを表す。図 10 では、図 1 と異なり、A-D 変換部 42 で、音声信号をデジタル変換されたデータが、認識応答制御部 46 に直接入力されるパスが用意されている。また図 11 は、利用者と認識応答装置の対話を示したものである。図 12 は、予め登録されている応答音声の一部を示したものである。

【0067】利用者と認識応答装置の対話は、図 11 に示すように、装置から利用者に対して、まずステップ S 11-1 で入力して項目に関する問合わせの応答音声

【0068】次に動作について説明する。図 11 において、ステップ S 11-2 で入力された音声の認識ができなかった時の対応方法を図 13 で説明する。まず図 11 のステップ S 11-1 に相当する問合わせのための応答音声を出力する（ステップ S 13-1）。次に利用者の入力音声の録音を開始（ステップ S 13-2）し、利用者の入力音声を認識する（ステップ S 13-3）。利用者の音声入力が終了すれば、入力音声の録音を終了する（ステップ S 13-4）。入力音声の録音は、図 10 では、利用者の電話 1 から入力された音声は、電話回線網 2、電話 I/F 部 41、A-D 変換部 42 を通して、認識応答制御部 46 から、システム制御部 6 に伝えられ、応答音声データ記憶部 7 で一つの応答音声ファイルとして、保持される。入力音声の認識は、A-D 変換部 42 から音声認識部 43 に音声信号が送られ、音声認識され、認識結果が、システム制御部 6 に伝えられる。

【0069】認識の結果が、「入力音声を認識できなかった」、すなわち図 13 のステップ S 13-5 における入力音声を認識したかという判断で、認識できなかった

14

場合、録音された入力音声の応答音声ファイルを応答音声データ記憶部 7 から取出し（ステップ S 13-6）、引き続き図 12 に示す 1203 の応答音声ファイルを取り出し、問い直しのための応答音声として、「（録音した音声）は認識できませんでした。もう一度言って下さい」という応答音声を作成し（ステップ S 12-7）、認識応答制御部 46、音声応答部 45 を通して出力する（ステップ S 12-8）。利用者は、自分の発声した音声を聞くことで、もし発声に誤りがあれば、次は正しく発声して音声入力することで、ステップ S 13-5 の入力音声の認識が正しくできたとして、次の対話に進む。

【0070】また別の対応方法を図 14 に示す。図 14 では、図 13 と異なり、ステップ S 14-7 で、問い直しのための応答音声として、利用者からの入力音声を録音した音声データを全て使うのではなく、音声の始端と終端の間の音声データを使っている。

【0071】ステップ S 14-3 における入力音声の認識は図 10 の音声認識部 43 で行われる。音声認識の前に音声認識部 43 では、音声データから音声の始端と終端を決め、認識応答制御部 46 に送る。認識応答制御部 46 では、A-D 変換部 42 から送られた利用者の音声に対して、音声認識部 43 から送られた始端、終端によって、音声データを切り出し、システム制御部 6 に伝えられ、応答音声データ記憶部 7 で、一つの応答音声ファイルとして、保持される。

【0072】認識の結果が、「入力音声を認識できなかった」、すなわち図 14 のステップ S 14-5 における入力音声を認識したかという判断で、認識できなかった場合、始端と終端ではさまれ録音された入力音声の応答音声ファイルを応答音声データ記憶部 7 から取出し（ステップ S 14-6）、引き続き図 12 に示す 1203 の応答音声ファイルを取り出し、問い直しのための応答音声として、「（録音した始端と終端の間の音声）は認識できませんでした。もう一度言って下さい」という応答音声を作成し（ステップ S 14-7）、問い直しのための応答音声を出力する（ステップ S 14-8）。利用者は、自分の発声した音声の内、認識応答装置で音声として捉えられた音声を聞くことで、もし発声に誤りがあれば、次は正しく発声し、また発声の始端や終端が曖昧であれば、次は明瞭に発声して音声入力することで、ステップ 14-5 における入力音声の認識が正しくでき、次の対話に進む。

【0073】次に図 11 において、ステップ S 11-2 で入力された音声の認識結果、複数の単語が出力された場合の対応方法を図 15 で説明する。図 10 で音声認識部 43 の認識結果が、一つの単語を出力すれば、確認のための応答音声を作成し（ステップ S 15-5）、音声応答部 45 を通して出力し（ステップ S 15-6）、これに対して入力された利用者の音声を認識する（ステップ S 15-7）。一方、音声認識部 43 が複数の単語を

出力した時、例えば、都市名を認識した結果、「富士（フジ）」、「宇治（ウジ）」、「久慈（クジ）」という3つの結果が得られ、入力音声として、これらの内の一つに特定できず、認識応答制御部46を通してシステム制御部7に報告された時（ステップS15-1）、システム制御部6は「いま言った内容に対応した番号をプッシュボタンで入力して下さい。1番富士、2番宇治、3番久慈。」という問合わせの応答音声を作成し（ステップS15-2）、音声応答部45を通して出力する（ステップS15-3）。

【0074】利用者は、応答音声を聞き、プッシュボタン入力し、認識応答装置は、プッシュボタン認識し、音声入力を確定させる（ステップS15-4）。

【0075】プッシュボタン認識は、図10でPB認識部44で利用者電話1からのプッシュボタン信号を認識し、結果がシステム制御部7に伝えられる。

【0076】また別の対応方法を図16に示す。図16では、図15と異なり、入力音声の確定方法として、プッシュボタンの認識に代わり、数字音声による認識（ステップS16-4）を使用する。

【0077】次に、言葉として似ている場合で、認識結果がひとつしかでず、しかもその結果が誤っている場合、例えば、利用者が「富士」と音声入力しているにもかかわらず、認識結果として、「宇治」だけが報告され、「富士」の音声入力が困難な場合の対応方法を、説明する。

【0078】認識応答装置は利用者からの音声入力を受け付ける前に予め、発声の似た言葉を図17のようにグループ化しておく。例えば「富士」「宇治」「久慈」という言葉は類似語として、ひとつのグループとしておく。

【0079】処理の流れを図18で説明する。利用者の音声認識した結果、ステップS18-1において、グループの内のひとつであった場合、そのグループに属する単語を取り出す（ステップS18-2）。例えば「富士」が認識結果として報告されれば、「宇治」や「久慈」を認識結果として報告の有無に関わらず、取り出す。次に、取り出した単語と、プッシュボタン入力案内の応答音声を作成する（ステップS18-3）。例えば、「いま言った内容に対応した番号をプッシュボタンで入力して下さい。1番富士、2番宇治、3番久慈。」という問合わせの応答音声を作成し、音声応答部45を通して出力する（ステップS18-4）。

【0080】利用者は、応答音声を聞き、プッシュボタン入力し、認識応答装置は、プッシュボタン認識し、音声入力を確定させる（ステップS18-5）。

【0081】プッシュボタン認識は、図10でPB認識部44で利用者電話1からのプッシュボタン信号を認識し、結果がシステム制御部7に伝えられる。

【0082】また別の対応方法を図19に示す。図19

では、図18と異なり、入力音声の確定方法として、プッシュボタンの認識に代わり、数字音声による認識（ステップS19-5）を使用する。

【0083】ここで、実施例2に関する各動作を、以下に要約して述べる。

【0084】音声認識部で、聞き違い易い複数の認識結果を出した場合、システム制御部にその結果を伝え、システム制御部では、応答音声として、複数の認識結果を利用者に伝え、プッシュボタンにより入力音声の確定を促す。

【0085】音声認識部で、聞き違い易い複数の認識結果を出した場合、システム制御部にその結果を伝え、システム制御部では、応答音声として、複数の認識結果を利用者に伝え、数字音声の入力により入力音声の確定を促す。

【0086】予め聞き違い易い音声をグループ化し、類似語グループとし、音声認識部が、類似語グループに属する音声の一つを結果として出した場合、システム制御部では、類似語グループに属する全ての音声を利用者に伝え、プッシュボタンにより入力音声の確定を促す。

【0087】予め聞き違い易い音声をグループ化し、類似語グループとし、音声認識部が、類似語グループに属する音声の一つを結果として出した場合、システム制御部では、類似語グループに属する全ての音声を利用者に伝え、数字音声の入力により入力音声の確定を促す。

【0088】実施例3。図20は、本発明の第3実施例に係る音声認識応答装置の全体構成を示すブロック図である。図20で、図1と同じ番号は、同じものを表す。図20では、電話I/F部41、A-D変換部42、音声認識部43、PB認識部44、音声応答部45、認識応答制御部46を合わせた部分が4で認識応答ブロック#1である。4と同じ構造をしたものが5の認識応答ブロック#2である。5の中には、電話I/F部51、A-D変換部52、音声認識部53、PB認識部54、音声応答部55、及び認識応答制御部56が含まれている。8はLAN、9はオペレータの端末でLAN9を介して、認識応答装置3のシステム制御部6とつながっている。10はオペレータの電話である。

【0089】図21は、処理の流れを示す図である。

【0090】図22は、図3の利用者電話のプッシュボタンについて、応答音声出力中にボタンプッシュすることで、オペレータの呼び出しができるボタンの位置を示したものである。

【0091】図23は、利用者と認識応答装置の対話が、電話発信、電話受信、案内メッセージの後は、認識応答装置からの問合わせ及び、利用者の回答を、[QA1]以下[QAブロック]として、対話が進んでいくことを示している。

【0092】図24は、オペレータが呼び出された時、オペレータの端末に表示される画面を示したものであ

る。

【0093】次に動作について説明する。図23に示すように、利用者は認識応答装置に対して、電話発信し、認識応答装置は電話受信後、案内メッセージを応答音声として出力する。案内メッセージに引き続き、[QA1]以下の各[QAブロック]の対話に入る。

【0094】対話にそった処理の流れを図21で説明する。まずステップS21-1で利用者からの電話接続をし、案内メッセージを出力する。次にステップS21-2で、問合わせのための応答音声を出し、ステップS21-3で利用者からの入力音声を認識し、認識結果からステップS21-4で「〇〇〇〇ですか?」と確認を求める応答音声を出し、ステップS21-5で利用者からの入力音声を認識し、ステップS21-6で認識結果が「はい」であれば、[QA1]を終了し、ステップS21-10で終了かどうか、すなわち問合わせ項目がないかどうかを判断し、なければステップS21-11で、終了または結果の応答音声を出して終わる。ステップS21-10で終了でなければ、ステップS21-2に戻り、次の[QA2]の問合わせのための応答音声の出力をする。ステップS21-6での認識結果が「はい」でなければ、ステップS21-7で再問合わせのための応答音声を出し、ステップS21-8でオペレータ呼び出しのプッシュボタン入力があれば、ステップS21-3で利用者からの音声入力を待って、その入力音声の認識をする。

【0095】図21における、ステップS21-2、ステップS21-4、ステップS21-7、ステップS21-11の応答音声の出力について、図20に示されたものの構成との関係に以下に述べる。図20において、システム制御部6が、応答音声データ記憶部7から応答音声データを引き出し、編集して、認識応答制御部46を通じて、音声応答部45に応答音声データをファイルとして送る。音声応答部45では、応答音声データファイルを分解した音声データを、A-D変換部42に送り、応答音声電話I/F部41を通じ、電話回線網2を通過して、利用者の電話1に送られる。

【0096】図21におけるステップS21-3、ステップS21-5の入力音声の認識につき図20に示されたものの構成との関係に以下に述べる。図20では、利用者の電話1から利用者が発声した音声、電話回線網2、電話I/F部41、A-D変換部42を通じて、音声信号として音声認識部43に伝えられる。音声認識部43では、伝えられた音声信号を音声認識し、認識結果を認識応答制御部46を通して、システム制御部6に伝える。

【0097】図23で[QA1]から順に利用者は認識応答装置と対話を進め、[QAn]の対話を終え、次の[QAn+1]の対話に入り、対話がうまく終了しなかったとする。すなわち図21で、ステップS21-5の

入力音声の認識の結果、ステップS21-6の認識結果の判断が繰り返し「いいえ」になり、ステップS21-3に帰るループから抜き出せなかったとする。または、同じ問合わせについて、一定回数以上の「いいえ」が続けば、「0」をボタンプッシュすれば、オペレータに接続される旨の応答音声をステップS21-7で出力する。

【0098】利用者は、この時図22に示すプッシュボタンのうち「0」を押してオペレータを呼び出す。

【0099】図20で、利用者の電話1から入力されたプッシュボタン信号は、電話回線網2、電話I/F部41、A-D変換部42を通じて、PB認識部44に伝えられる。PB認識部44では、プッシュボタン信号を認識し、プッシュボタンにより「0」が入力されたことを、認識制御部46を通じて、システム制御部6に伝えられる。システム制御部6は、認識制御部56を通じて電話I/F部51からオペレータの電話10に電話発信をする。オペレータは、電話を受信し、オペレータ端末9を起動すると、システム制御部6からLAN8を通じて、オペレータ端末9には、図24の画面が表示される。

【0100】オペレータは、図24の画面を見て、[QAn+1]の対話を認識応答装置を通じて行う。利用者とオペレータの対話は、4の認識応答ブロック#1と、5の認識応答ブロック#2が、システム制御部6の制御をもとに、音声信号を受け渡しすることで行う。すなわち、4の認識応答ブロック#1で電話I/F部41を通して入力された、利用者の音声は、A-D変換部42、認識応答制御部46を通して、システム制御部6の制御で、5の認識応答ブロック#2の認識応答制御部56に伝わり、電話I/F部を通して、オペレータの電話10に伝わる。オペレータの音声は、上記とは反対の経路で利用者の電話1に伝えられる。

【0101】オペレータは利用者の対話の結果、[QAn+1]の利用者の回答を、オペレータ端末9から入力し、入力データはLAN9を通じてシステム制御部6に伝えられる。システム制御部6は利用者とオペレータとの対話がどの段階まで行われたかを認識することになる。利用者との対話が正常になったとオペレータが判断すれば、オペレータ端末9から、認識応答装置の対話を指示し、オペレータは受話器を置く。

【0102】システム制御部6は、オペレータ端末9からの指示により、認識応答装置と利用者との対話を再開する。この再開は、オペレータによる対話の続きからなされる。すなわち、オペレータにより、オペレータ端末9から入力され、図24で[QAn+3]まで回答が入力されていれば、このことをシステム制御部は、認識しているため、[QAn+4]から利用者との対話が再開される。

【0103】この実施例では、利用者がオペレータとの

対話を希望した場合、プッシュボタン操作により、オペレータに接続され、利用者とオペレータの対話ができ、その後、オペレータが、利用者と装置の対話が可能と判断したら、オペレータからの端末操作により、再び、利用者と装置の対話を行うことができる。

【0104】

【発明の効果】以上詳細に説明してきたように、本発明によれば、利用者からの指示により、応答音声の出力の態様が変わるので、利用者にとって自分にあった応答音を聞くことができる。

【0105】また、本発明によれば、利用者の指示により、応答音声の早送りや巻き戻しが簡単にできるので、装置に慣れた利用者が応答音を聞き飛ばしたり、装置に不慣れた利用者が繰り返し応答音を聞いたりすることができ、利用者の要求に応じて装置を効率よく利用することができる。

【0106】また、本発明によれば、利用者の指示により、簡単に装置からの応答音声による質問をスキップして一つ先の質問に移れるので、利用者は質問への回答を保留して、次の質問への入力ができ、利用者にとって、

考える順序を柔軟にすることができる。

【0107】また、本発明によれば、利用者の指示により、簡単に装置からの応答音声による質問を戻して、一つ前の質問に移れるので、利用者は誤った入力あるいは、装置の誤認識により利用者の意図しない方向は対話が進んだ場合、元の正しい対話まで戻ることができる。

【0108】また、本発明によれば、音声認識ができなかった場合、利用者の発声を録音し、利用者に聞かせることで、音声認識できなかった原因が、利用者の誤発声にあった場合、利用者に注意を促し、次の認識が適正に

できるようにする。

【0109】また、本発明によれば、音声認識ができなかった場合、利用者の発声を録音し、装置が音声の始端と終端ではさまれた部分を利用者に聞かせることで、音声認識できなかった原因が、利用者の発声の始まりあるいは終わりの曖昧にあった場合、利用者に明瞭に発声するよう注意を促し、次の認識がうまくできるようになる。

【0110】また、本発明によれば、複数の音声認識の結果がでた場合、利用者からの応答入力により、発声の似た言葉について、利用者からの音声入力、装置の音声認識を繰り返すことなく、簡単に入力音声を確定させることができる。

【0111】また、本発明によれば、発声の似た言葉について、予めグループ化しておき、グループ内の一つが、音声認識の結果としてでた場合、グループ内にある発声の似た言葉を利用者に案内し、利用者の応答入力に基き、入力音声を確定させる。これによって、装置の誤認識により、繰り返し音声入力しても、目的の言葉を認識できない場合、簡単に目的の言葉を認識させることが

できる。

【0112】また、本発明によれば、利用者が設定した応答音声の速度、大きさ、男性あるいは女性の声など、応答音声に関する値を記憶しておき、次回同じ利用者から装置の利用があったとき、記憶結果をもとに、利用者にあわせた応答音声を提供できる。

【0113】また、本発明によれば、利用者と装置の対話がうまく進まなかった時、利用者からの指示によりオペレータを呼び出し、その後正常な対話に戻った時に、再び利用者と装置の対話をするようにしたので、オペレータが利用者と対話する時間を少なくすることができ、オペレータが効率良く作業を進められる。

【図面の簡単な説明】

【図1】 この発明の第1実施例に係る音声認識応答装置の全体構成を示すブロック図である。

【図2】 この発明の第1実施例に係るプッシュボタン入力に対するフロー図である。

【図3】 図中1の利用者電話のプッシュボタンを示した図である。

【図4】 この発明の第1実施例に係る図中1の利用者電話のプッシュボタンについて、働く機能を示した図である。

【図5】 この発明の第1実施例に係る利用者と音声認識応答装置の対話における問合わせのための応答音声と応答音声ファイル番号の対応を示した図である。

【図6】 この発明の第1実施例に係るプッシュボタン「7」の入力と音声出力を示した図である。

【図7】 この発明の第1実施例に係るプッシュボタン「9」の入力と音声出力を示した図である。

【図8】 この発明の第1実施例に係るプッシュボタン「*」の入力と音声出力を示した図である。

【図9】 この発明の第1実施例に係るプッシュボタン「#」の入力と音声出力を示した図である。

【図10】 この発明の第2実施例に係る音声認識応答装置の全体構成を示すブロック図である。

【図11】 この発明の第2実施例に係る利用者と音声認識応答装置の対話を示す図である。

【図12】 この発明の第2実施例に係る予め登録されている応答音声の一部を示した図である。

【図13】 この発明の第2実施例に係る入力された音声認識できなかった時の対応方法を示したフロー図である。

【図14】 この発明の第2実施例に係る入力された音声認識できなかった時の別の対応方法を示したフロー図である。

【図15】 この発明の第2実施例に係る入力された音声認識の結果、複数の単語がでた時の対応方法を示したフロー図である。

【図16】 この発明の第2実施例に係る入力された音声認識の結果、複数の単語がでた時の別の対応方法を

21

示したフロー図である。

【図 17】 この発明の第 2 実施例に係る発声の似た単語のグループ化を示した図である。

【図 18】 この発明の第 2 実施例に係る発声の似た単語が認識結果として得られた場合の対応方法を示した図である。

【図 19】 この発明の第 2 実施例に係る発声の似た単語が認識結果として得られた場合の別の対応方法を示した図である。

【図 20】 この発明の第 3 実施例に係る音声認識応答装置の全体構成を示すブロック図である。

【図 21】 この発明の第 3 実施例に係る誤認識が多い場合の対応方法を示したフロー図である。

【図 22】 この発明の第 3 実施例に係る図中 1 の利用者電話のプッシュボタンについて、働く機能を示した図である。

【図 23】 この発明の第 3 実施例に係る利用者と認識応答装置の対話を示す図である。

【図 24】 この発明の第 3 実施例に係る図中 9 のオペ

22

レータ端末への表示を示す図である。

【図 25】 特定語を認識することで、聞き逃した音声再度聞くようにした従来技術のブロック図である。

【図 26】 特定語を認識することで、聞き逃した音声再度聞くようにした従来技術の処理フロー図である。

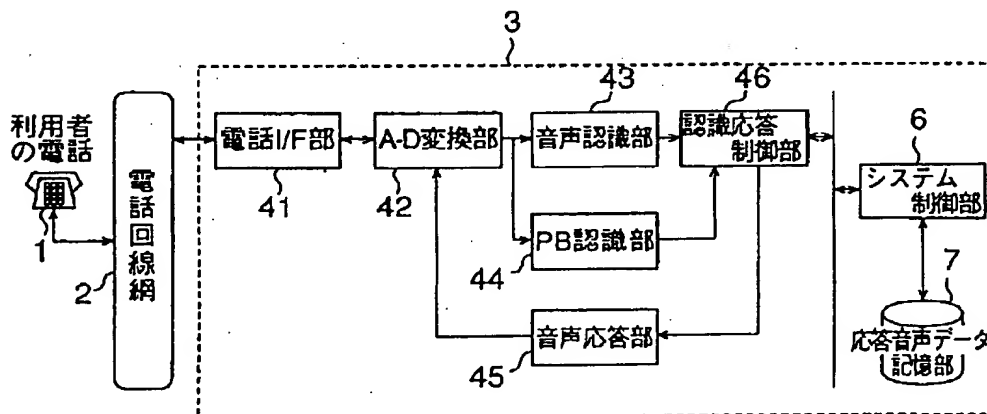
【図 27】 音声認識が不能になった場合オペレータに切り替えるようにした従来技術のブロック図である。

【図 28】 音声認識が不能になった場合オペレータに切り替えるようにした従来技術の処理フロー図である。

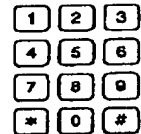
【符号の説明】

1 利用者の電話、2 電話回線網、3 音声認識応答装置、4 認識応答ブロック #1、5 認識応答ブロック #2、6 システム制御部、7 応答音声データ記憶部、8 LAN、9 オペレータの端末、10 オペレータの電話、41、51 電話 I/F 部、42、52 A-D 変換部、43、53 音声認識部、44、54 PB 認識部、45、55 音声応答部、46、56 認識応答制御部。

【図 1】



【図 3】

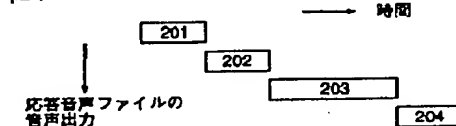


【図 5】

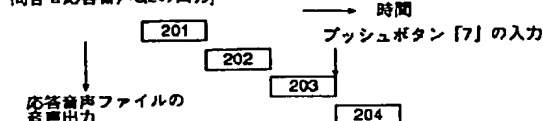
問合わせのための応答音声	応答音声ファイル番号
Q1	101,102,103
Q2	201,202,203,204
Q3	301,302,303,304,305
Q4	401,402,403,404

【図 6】

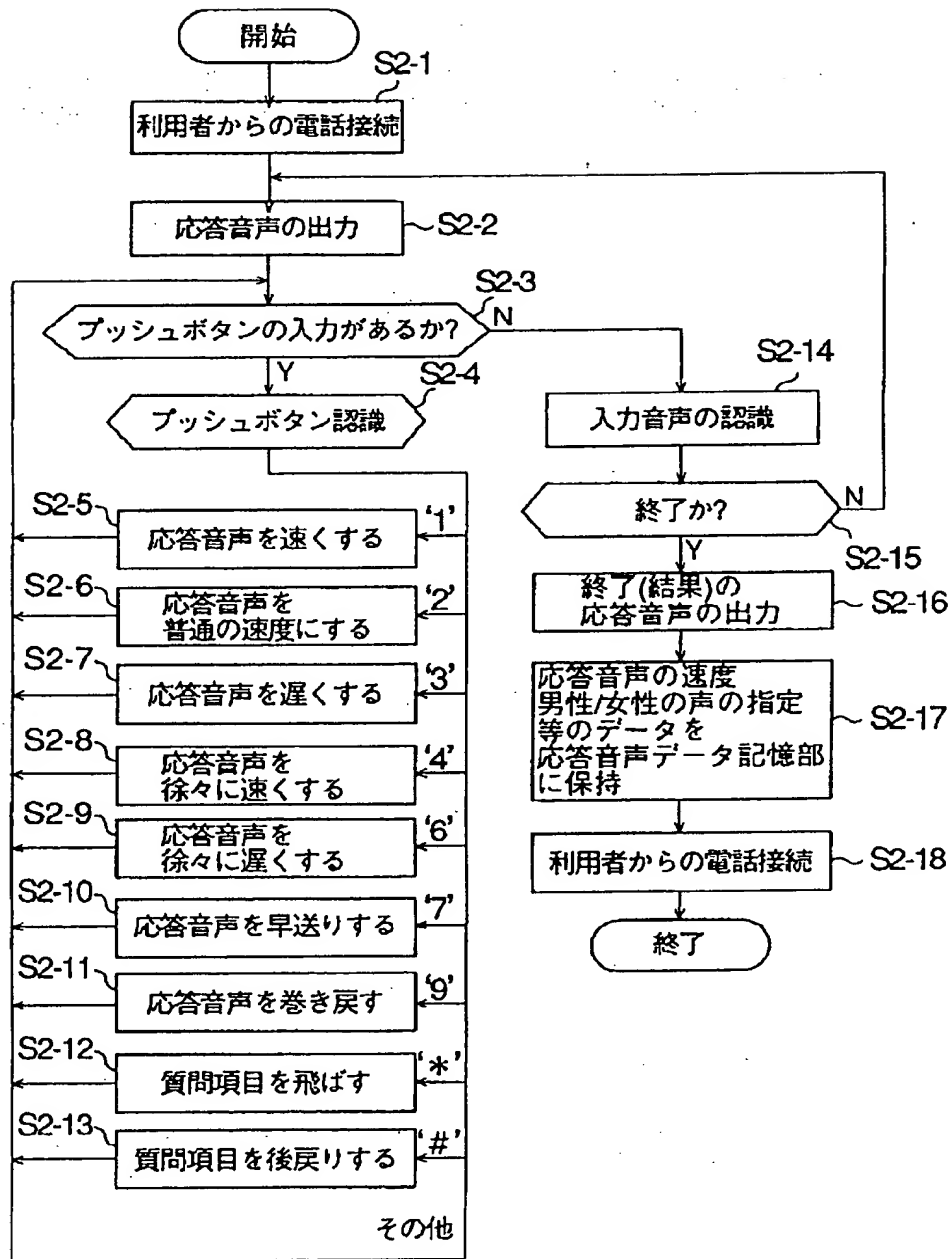
【通常の問合わせ応答音声 Q2 の出力】



【プッシュボタン「7」入力と問合わせ応答音声 Q2 の出力】



【図 2】



【図 1 1】

【音声認識応答装置】	【利用者】
S11-1 「.....を言って下さい」	S11-2 「0000」
S11-3 「0000ですか」	S11-4 「はい」

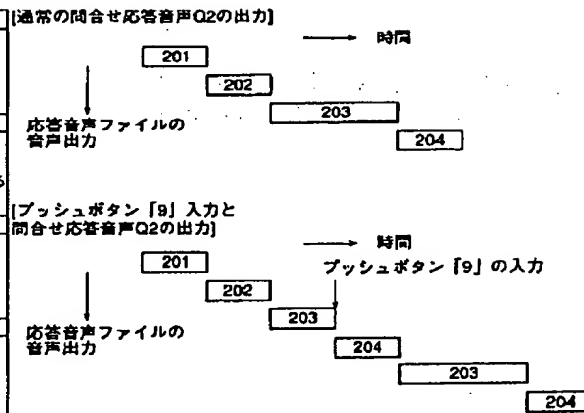
【図 1 2】

応答音声 ファイル番号	応答音声の内容
.....
1201	「を、言ってください」
1202	「ですか。」
1203	「は、認識できませんでした。もう一度言って下さい。」

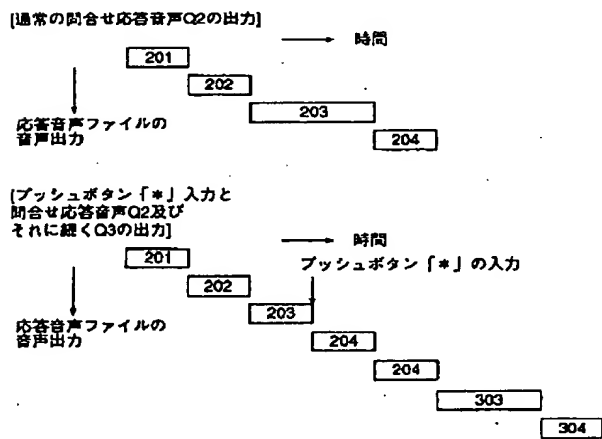
【図 4】

1	2	3
速い速度の 応答音声	普通の速度の 応答音声	遅い速度の 応答音声
4	5	6
応答音声を 徐々に速くする		応答音声を 徐々に遅くする
7	8	9
応答音声の 早送り		応答音声の 巻戻し
*	0	#
応答音声の 飛ばし		応答音声の 繰り返し

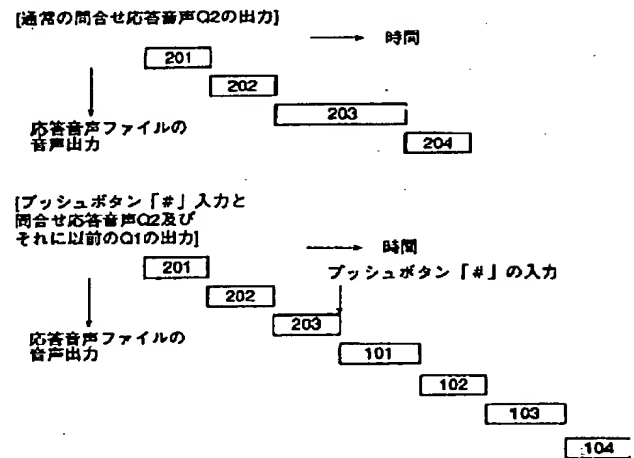
【図 7】



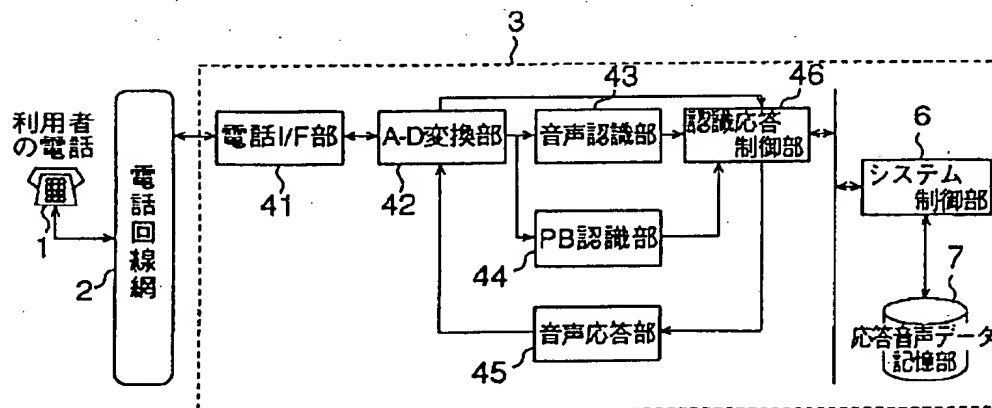
【図 8】



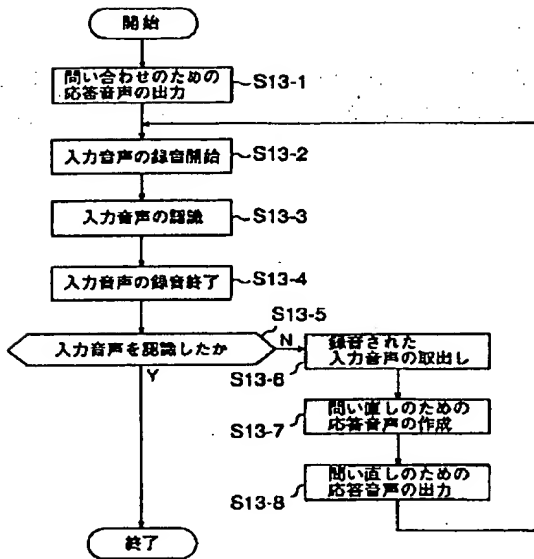
【図 9】



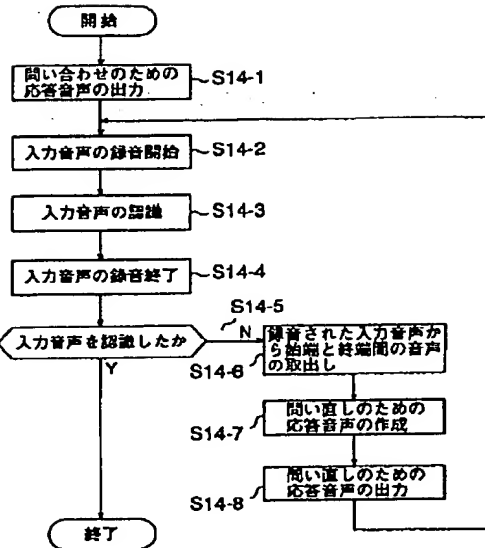
【図 10】



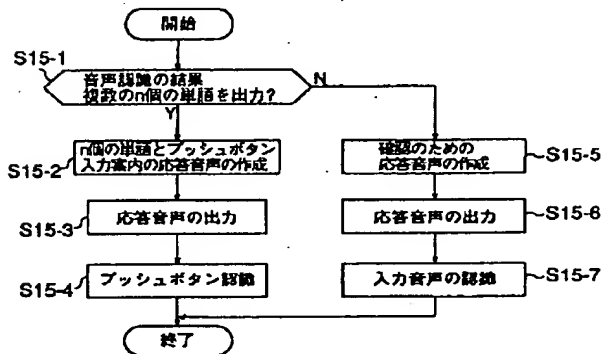
【図 13】



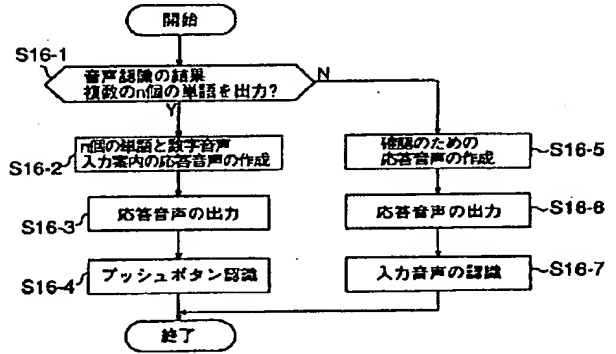
【図 14】



【図 15】



【図 16】



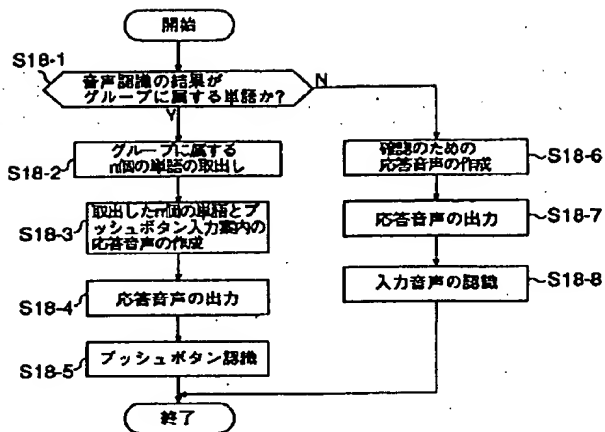
【図 17】

グループ番号	発声の似た単語
.....
1701	富士(フジ)、宇治(ウジ)、久慈(クジ)
1702	須賀川(スカガワ)、深川(フカガワ)
1703	安城(アンジョウ)、三島(サンジョウ)
1704	佐賀(サガ)、加賀(カガ)
1705	秋川(アキカワ)、滝川(タキカワ)
.....

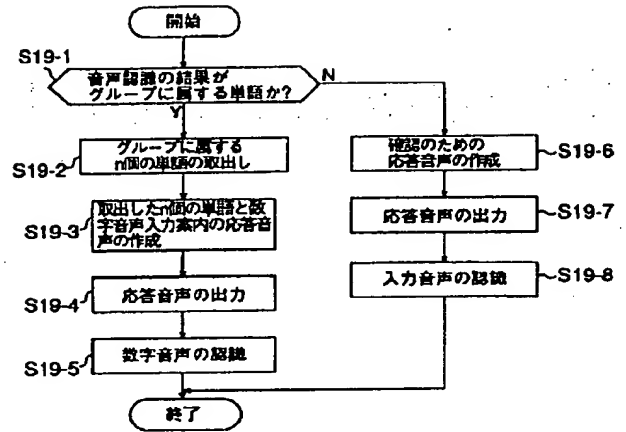
【図 22】

1	2	3
4	5	6
7	8	9
*	0	#
オペレータの 呼出し		

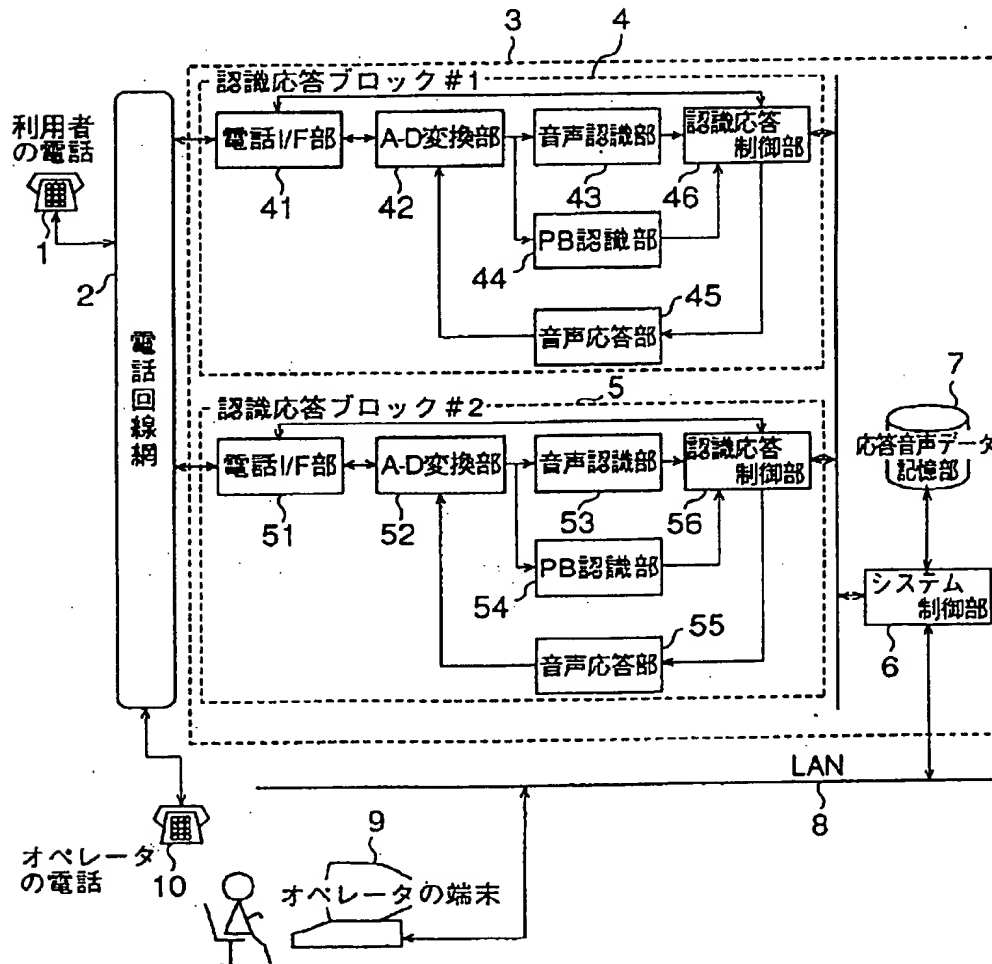
【図18】



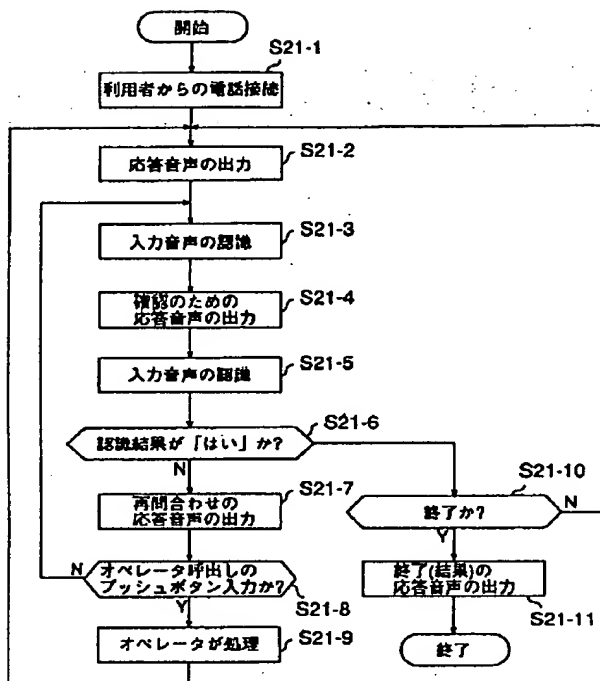
【図19】



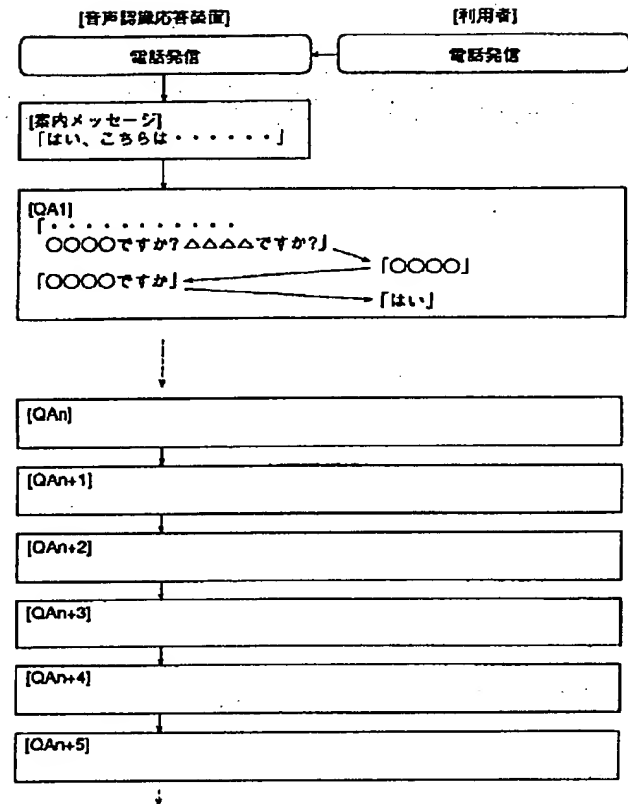
【図20】



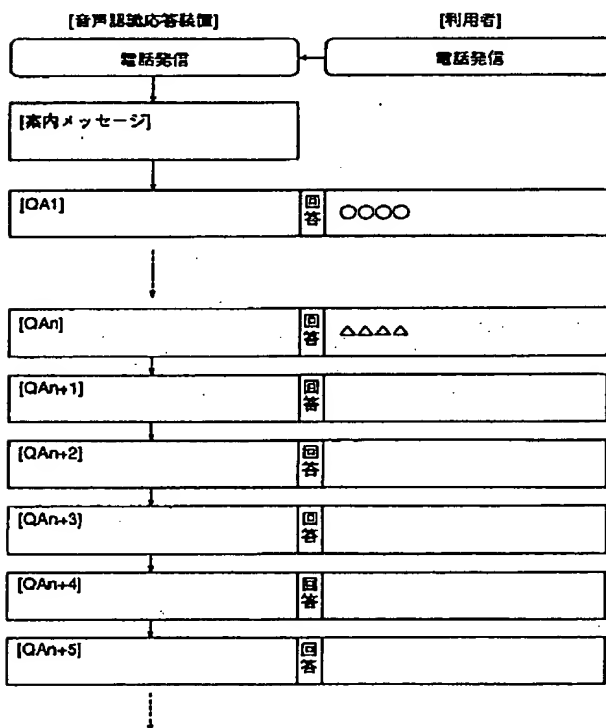
【図 2 1】



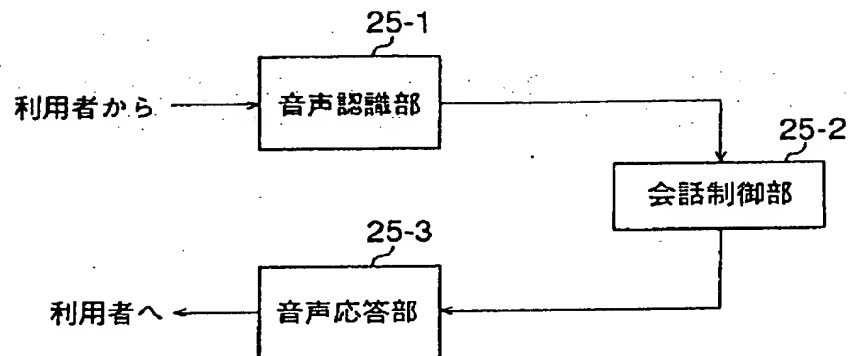
【図 2 3】



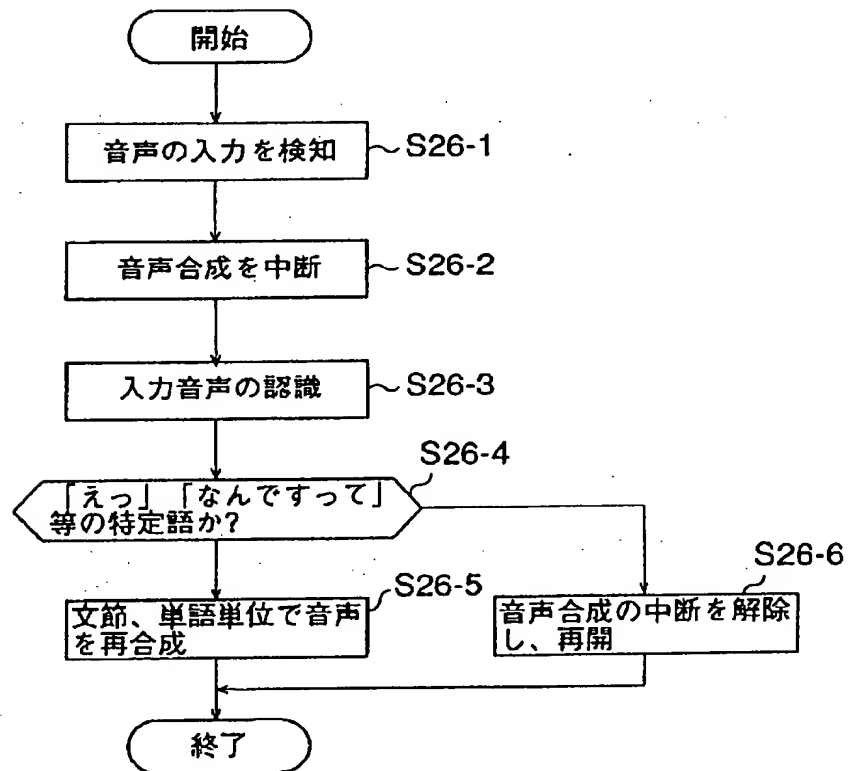
【図 2 4】



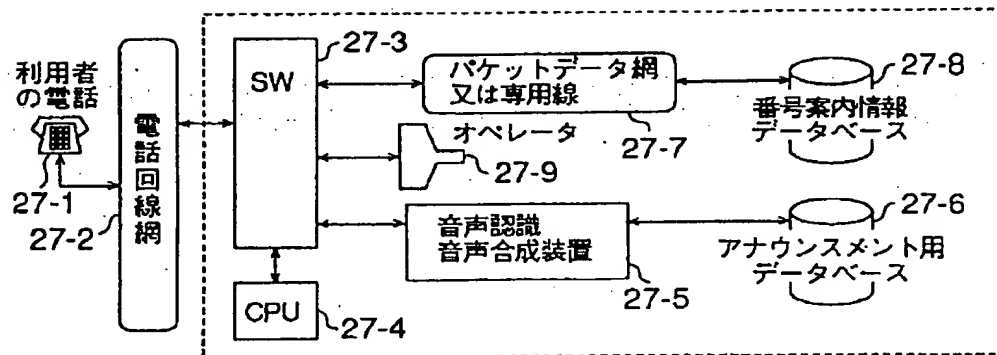
【図25】



【図26】



【図 27】



【図 28】

